REMARKS

This application contains claims 1-3, 5-7, 9-18, 21-24, 26-31 and 33-35. Claims 1, 15, 22, 27, 29 and 34 are hereby amended. No new matter has been added. Reconsideration is respectfully requested.

Applicant thanks Examiners Spooner and Storm for the courtesy of an interview granted to Applicant's representative, Daniel Kligler (Reg. No. 41,120), at the USPTO on March 17, 2005. At the interview, Applicant's representative pointed out the distinction of the present invention over Halstead Jr., et al. (U.S. Patent 5,963,893), specifically in terms of evaluating and using the frequency of occurrence of the patterns of words. It was agreed that if Applicant were to amend the independent claims to recite that the frequencies of the patterns of the words are evaluated independently of the lemmas of the words, the claims would then distinguish over Halstead. Examiner Storm also pointed out a number of informalities in claim 1 and requested that they be corrected. Applicant has amended the claims accordingly.

Claim 1 was objected to because in the previous amendment in this case, the claim was erroneously marked "Original," rather than "Currently amended." The claim has been properly marked in the present amendment.

Claims 1, 9-13, 22, 26, 29 and 33 were rejected under 35 U.S.C. 102(e) over Halstead, while claims 2, 3, 23, 24, 30 and 31 were rejected under 35 U.S.C. 103(a) over Halstead. Applicant has amended independent claims 1, 22 and 29, as agreed in the interview, in order to further clarify the distinction of the present invention over the cited art and to correct a number of informalities. These claims recite a method, software product and apparatus for morphological disambiguation of an input string. The claims have been amended to state that in evaluating analyses of the string, the relative frequency of occurrence of the pattern in each of the analyses is determined <u>independently of the lemma to which the pattern is applied</u>. This added limitation is supported literally in the specification on page 5, lines 14-19. The "lemma" of a word is defined in the specification (page 2, lines 5-8) as the lexical base form of the word, i.e., the root, or stem, of the word, independent of the linguistic pattern.

Halstead describes a word breaking facility for identifying words within a Japanese text string based on morphological processing (abstract). In one aspect of this facility, n-gram templates are provided that specify a pattern of character types found in stems and the likelihood of the template occurring in a stem (col. 1, line 64 - col. 2, line 1). Morphological analysis is performed on an input string of Japanese characters to give a directed acyclic graph. Each path through the graph represents a morphological analysis of at least a portion of the input string. The paths are scored to favor paths that include analysis for a greater portion of the input string and paths that include a greater number of bound morphemes. The highest-scoring path is selected and applied as the morphological analysis of the input string (col. 2, lines 13-27).

In maintaining the rejection of these claims in the present Official Action, the Examiner held that Halstead teaches finding patterns of words and determining a relative frequency of occurrence of the patterns. In this regard, the Examiner cites cols. 7, 8 and 9 in Halstead. In col. 7 (lines 5-26), Halstead examines "paradigms" in order to determine which characters may be strung together into a word in accordance with grammatical rules. Even if these paradigms are considered to be analogous to the patterns of claim 1, however, Halstead is not concerned with the frequency of the paradigms. Rather, Halstead scores the paths in his morphological map depending on the number of characters in the path, the depth of the analysis, and the morpheme bigram probabilities (col. 8, lines 35-39). None of these factors has anything to do with the frequency of a "pattern" (as defined in claim 1) that might be associated with a given morphological path. Rather, Halstead's scoring "emphasizes analyses that have a larger number of characters and a deeper number of levels." Thus, Halstead's scoring is intimately dependent on the specific characters in the word (i.e., the lemma), rather than the linguistic pattern.

Hence, when Halstead states that "the highest scoring paths are kept" (col. 9, line 8), he means that the paths are chosen based on the number of characters that can be combined and the likelihood that the given combination of characters correctly forms a word. Halstead gives no regard to whether the pattern of that word is frequent or infrequent. Amended claims 1, 22 and 29 are therefore believed to be patentable over Halstead. In view of the patentability of these independent

claims, dependent claims 2, 3, 9-13, 23, 24, 26, 30, 31 and 33 are also believed to be patentable.

Claims 5-7 and 14 were rejected under 35 U.S.C. 103(a) over Halstead in view of Zamora (U.S. Patent 4,862,408). These claims depend from claim 1. In view of the patentability of claim 1 as amended, claims 5-7 and 14 are believed to be patentable, as well.

Claims 15-18, 21, 27, 28, 34 and 35 were rejected under 35 U.S.C. 103(a) over Zamora in view of Halstead. Applicant has amended independent claims 15, 27 and 34 in a manner similar to the amendment of claims 1, 22 and 29, so as to clarify the distinction of the present invention over the cited art.

In rejecting claims 15, 27 and 34, the Examiner referred to the same passages in Halstead as were cited against claims 1, 22 and 29. Claims 15, 27 and 34, as amended, recite a method, apparatus and software product for searching a corpus of text, wherein candidate analyses of words in the corpus are selected based on the relative frequency of occurrence of their respective patterns independent of the lemmas to which the patterns are applied. The lemmas of the selected analyses are entered in an index of the corpus, to which a search query may then be applied. Amended claims 15, 27 and 34 are thus believed to distinguish over Halstead for the reasons stated above, as agreed in the interview.

Zamora describes a method for analyzing text using a paradigm. He creates a file structure in which each word in a list of words (or "dictionary") is associated with a set of paradigm references. These references generate all forms of each of the lemmas of the words in the list (abstract). In other words, Zamora uses all possible linguistic forms of each of the words in a given list (col. 2, lines 66-68), without discriminating between the more and less frequent forms, as required by claims 15, 27 and 34.

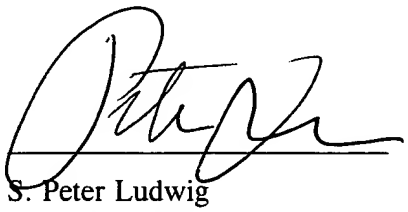
Thus, claims 15, 27 and 34, as amended, are believed to be patentable over Halstead and Zamora. In view of the patentability of these independent claims, dependent claims 16-18, 21, 28 and 35 are believed to be patentable, as well.

Applicant believes the amendments and remarks presented hereinabove to be fully responsive to all of the objections and grounds of rejection raised by the Examiner. In view of these amendments and remarks, Applicant respectfully

submits that all of the claims in the present application are in order for allowance. Notice to this effect is hereby requested.

Date: March 30, 2005

Respectfully submitted,



S. Peter Ludwig
Reg. No. 25,351
Attorney for Applicants

DARBY & DARBY, P.C.
P.O. Box 5257
New York, NY 10150-5257
212-527-770